DOCUMENT RESUME

ED 241 564 TM 840 088

AUTHOR Rock, Donald; Werts, Charles
TITLE An Analysis of Time-Related Score Increments and/or Decrements for GRE Repeaters across Ability and Sex Groups.
INSTITUTION Educational Testing Service, Princeton, N.J.; Graduate Record Examinations Board, Princeton, N.J.
REPORT NO GREB-77-9R
PUB DATE Mar 80
NOTE 52p.
AVAILABLE FROM Educational Testing Service, Publications Order Services, Dept. I-101, Princeton, NJ 08541.
PUB TYPE Reports - Research/Technical (143)

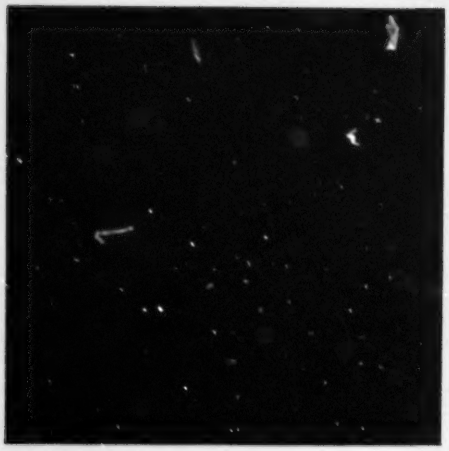
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Ability; *Achievement Gains; *College Entrance Examinations; Higher Education; Quantitative Tests; *Scores; *Sex Differences; Test Wiseness; Verbal Tests
IDENTIFIERS *Graduate Record Examinations; *Retesting

ABSTRACT

The purpose of this study was to obtain information on both the number of individuals who retest and their patterns of score gain (or decrement) by sex and ability. Individuals who retested only once were found to gain about 26-27 points on the Graduate Record Examination (GRE) verbal test and about 23 points on the GRE quantitative test. This finding is tempered by the fact that repeaters are a self-selected group whose average scores were below the mean on initial testing; average gains on retest by a random sample of all test takers almost certainly would be somewhat less. Individuals who retested more than once were, on the average, of lower ability than the single retest individuals. The multiple retesters gained less on the average than the higher ability single retest individuals. The amount of verbal gain was positively related to the length of time between the initial test and the first retest. This result was interpreted primarily as an increase in knowledge and to a lesser extent an increase in test-taking skills. There was no relationship between time lapse and quantitative gains. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED241564



GRE

GRADUATE RECORD EXAMINATIONS

AN ANALYSIS OF TIME-RELATED SCORE INCREMENTS AND/OR DECREMENTS FOR GRE REPEATERS ACROSS ABILITY AND SEX GROUPS

Donald Rock
and
Charles Werts

GRE Board Research Report GREB No. 77-9R

March 1980

This report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

EDUCATIONAL TESTING SERVICE, PRINCETON, NJ

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Weidenmiller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- ☑ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

GRE BOARD RESEARCH REPORTS
FOR GENERAL AUDIENCE

Altman, P. A. and Wallmark, M. M. A Summary of Data from the Graduate Programs and Admissions Manual. GRE Board Research Report GREB No. 74-1R, January 1975.

Baird, L. L. An Inventory of Documented Accomplishments. GRE Board Research Report GREB No. 77-3R, June 1979.

Baird, L. L. Cooperative Student Survey (The Graduates [$2.50 each], and Careers and Curricula). GRE Board Research Report GREB No. 70-4R, March 1973.

Burns, R. L. Graduate Admissions and Fellowship Selection Policies and Procedures (Part I and II). GRE Board Research Report GREB No. 69-5R, July 1970.

Centra, J. A. How Universities Evaluate Faculty Performance: A Survey of Department Heads. GRE Board Research Report GREB No. 75-5bR, July 1977. ($1.50 each)

Centra, J. A. Women, Men and the Doctorate. GRE Board Research Report GREB No. 71-10R, September 1974. ($3.50 each)

Clark, M. J. The Assessment of Quality in Ph.D. Programs: A Preliminary Report on Judgments by Graduate Deans. GRE Board Research Report GREB No. 72-7aR, October 1974.

Clark, M. J. Conditions Influencing the Career Accomplishments of Recent Ph.D.'s. GRE Board Research Report GREB No. 76-2R, June 1979.

Clark, M. J. Program Review Practices of University Departments. GRE Board Research Report GREB No. 75-5aR, July 1977. ($1.00 each)

Donlon, T. F. Annotated Bibliography of Test Speededness. June 1979.

Flaugher, R. L. The New Definitions of Test Fairness In Selection: Developments and Implications. GRE Board Research Report GREB No. 72-4R, May 1974.

Fortna, R. O. Annotated Bibliography of the Graduate Record Examinations. July 1979.

Frederiksen, N. and Ward, W. C. Measures for the Study of Creativity in Scientific Problem-Solving. May 1978.

Hartnett, R. T. The Information Needs of Prospective Graduate Students. GRE Board Research Report GREB No. 77-8R, October 1979.

Hartnett, R. T. and Willingham, W. W. The Criterion Problem: What Measure of Success in Graduate Education? GRE Board Research Report GREB No. 77-4R, March 1979.

Knapp, J. and Hamilton, I. B. The Effect of Nonstandard Undergraduate Assessment and Reporting Practices on the Graduate School Admissions Process. GRE Board Research Report GREB No. 76-14R, July 1978.

Lannholm, G. V. and Parry, M. E. Programs for Disadvantaged Students in Graduate Schools. GRE Board Research Report GREB No. 69-1R, January 1970.

Miller, R. and Wild, C. L. Restructuring the Graduate Record Examinations Aptitude Test. GRE Board Technical Report, June 1979.

Reilly, R. R. Critical Incidents of Graduate Student Performance. GRE Board Research Report GREB No. 70-5R, June 1974.

Rock, D., Werts, C. An Analysis of Time Related Score Increments and/or Decrements for GRE Repeaters across Ability and Sex Groups. GRE Research Report GREB No. 77-9R, April 1979.

Rock, D. A. The Prediction of Doctorate Attainment in Psychology, Mathematics and Chemistry. GRE Board Research Report GREB No. 69-6aR, June 1974.

Schrader, W. B. Admissions Test Scores as Predictors of Career Achievement in Psychology. GRE Board Research Report GREB No. 76-1R, September 1978.

Wild, C. L. Summary of Research on Restructuring the Graduate Record Examinations Aptitude Test. February 1979.

Wild, C. L. and Durso, R. Effect of Increased Test-Taking Time on Test Scores by Ethnic Group, Age, and Sex. GRE Board Research Report GREB No. 76-6R, June 1979.

Wilson, K. M. The GRE Cooperative Validity Studies Project. GRE Board Research Report GREB No. 75-8R, June 1979.

Wiltsey, R. G. Doctoral Use of Foreign Languages: A Survey. GRE Board Research Report GREB No. 70-14R, 1972. (Highlights $1.00, Part I $2.00, Part II $1.50).

Witkin, H. A.; Moore, C. A.; Oltman, P. K.; Goodenough, D. R.; Friedman, F.; and Owen, D. R. A Longitudinal Study of the Role of Cognitive Styles in Academic Evolution During the College Years. GRE Board Research Report GREB No. 76-10R, February 1977 ($5.00 each).

An Analysis of Time-related Score Increments and/or Decrements for GRE Repeaters Across Ability and Sex Groups

Donald Rock

and

Charles Werts

GRE Board Research Report GREB No. 77-9R

March 1980

Copyright © 1980 by Educational Testing Service. All rights reserved.

Contents

	Page
Abstract	i
List of Tables	ii
List of Figures	iii
List of Appendices	iv
Introduction	1
Procedure	2
Results and Discussion	4
Score Gains by Men and Women	8
Score Gains and Time Between Tests	13
Regression Effects and "True" Gain	15
Characteristics of Repeaters	17
Simplex Model	21
Summary and Conclusions	22
References	24
Appendices	25

Abstract

The purpose of this study was to obtain information on both the number of individuals who retest and their patterns of score gain (or decrement) by sex and ability. Individuals who retested only once were found to gain about 26 - 27 points on the verbal test and about 23 points on the quantitative test. This finding is tempered by the fact that repeaters are a self-selected group whose average scores were below the mean on initial testing; average gains on retest by a random sample of all test takers almost certainly would be somewhat less. Individuals who retested more than once were, on the average, of lower ability than the single retest individuals. The multiple retesters gained less on the average than the higher ability single retest individuals. The amount of verbal gain was positively related to the length of time between the initial test and the first retest. This result was interpreted primarily as an increase in knowledge and to a lesser extent an increase in test-taking skills. There was no relationship between time lapse and quantitative gains.

List of Tables

		Page
Table 1:	Mean and Total Gains for Verbal and Quantitative Scores by Administration Populations	3
Table 2:	Mean and Total Gains for Verbal and Quantitative Scores by Sex and Administration Populations	13
Table 3:	Regression of Retest Scores on Initial Scores and Time Lapse by Sex and Aptitude Areas	14
Table 4:	Regression of Repeater Group Membership on Demographic Variables	17
Table 5:	Regressions of Verbal and Quantitative Retest Scores on Initial Scores and Significant Demographic Variables . .	18
Table 6:	Mean Scores on Initial Testing and Mean Gains on First Retesting for Selected Subgroups	20

List of Figures

Page

Figure 1: Verbal mean score changes by number of repeated administrations 5

Figure 2: Quantitative mean score changes by number of repeated administrations 6

Figure 3: Verbal mean score changes for women by number of repeated administrations 9

Figure 4: Verbal mean score changes for men by number of repeated administrations 10

Figure 5: Quantitative mean score changes for women by number of repeated administrations 11

Figure 6: Quantitative mean score changes for men by number of repeated administrations 12

List of Appendices

Page

Appendix A: Verbal Score Changes on Retesting by Level of Initial Score (Single Repeaters Only). 25

Appendix B: Quantitative Score Changes on Retesting by Level of Initial Score (Single Repeaters Only) 27

Appendix C: Verbal Score Gains for Males by Time Lapse (Single Repeaters Only) 29

Appendix D: Verbal Score Gains for Females by Time Lapse (Single Repeaters Only) 30

Appendix E: Quantitative Score Gains for Males by Time Lapse (Single Repeaters Only) 31

Appendix F: Quantitative Score Gains for Females by Time Lapse (Single Repeaters Only) 32

Appendix G: Frequency Distribution of Verbal Gains on First Retesting 33

Appendix H: Frequency Distribution of Quantitative Gains on First Retesting 34

Appendix I: Background Information Questions 35

Appendix J: Intercorrelations, Means, and Standard Deviations for the Variables Reported in Table 4 and Table 5 37

Appendix K: Intercorrelations of Scores on Successive Retests for the Total Four-Administration Population . . 38

Appendix L: Intercorrelations of Verbal Scores on Successive Retests for the Four-Administration Population, by Sex 39

Appendix M: Intercorrelations of Quantitative Scores on Successive Retests for the Four-Administration Population, by Sex 40

An Analysis of Time-related Score Increments and/or Decrements for GRE Repeaters Across Ability and Sex Groups

Donald Rock and Charles Werts

Introduction

The purpose of this study was to obtain information on (a) how many persons repeat the Graduate Record Examinations (GRE) Aptitude Test one time or more; (b) whether there is a time-related score decrement (or increment) for those who repeat; (c) whether the pattern of increment or decrement differs by sex and/or ability area (ability area may be understood to mean the measured verbal and quantitative aptitude areas); and (d) whether one-time repeaters differ from multiple repeaters in ability and/or sex group membership.

The information provided by answers to the questions listed above should help candidates decide whether or not, or under what conditions, they should repeat the test, and should help institutions interpret multiple test scores when making admissions decisions.

In 1958 Levine and Angoff investigated the effect of repetition on the Scholastic Aptitude Test for high school juniors and seniors. They found an average gain of ten points for the first retest, an additional gain of ten points for a second retest, and no gain for a third retest.

A number of other studies investigating the effects of repeating a test have been reviewed by Weiss (1961). Weiss concluded that (a) practice improves performance on tests of intelligence and scholastic aptitude; (b) the effects of a first and second practice for mental tests are significant, but thereafter the effect diminishes; (c) practice effects vary with the time interval between practice and final tests (significant practice effects are demonstrated for time intervals of two weeks to three months); and (d) practice effects interact with mental ability. The more intelligent test takers appear to benefit the most.

Campbell, Hilton, and Pitcher (1967) sent a questionnaire to those individuals who took tests in both the November 1962 and the January 1963 administrations of the Graduate Record Examinations. This questionnaire was designed to obtain such information as the students' reasons for taking the tests when they did, the amount of

effort applied during each administration, and the relevance of courses taken during the interim. The questionnaire did not specifically ask if they attended a "coaching" school. Campbell et al. found an increase of 25 points on the verbal section and 16 points on the quantitative section. For those individuals who repeated an Advanced Test, the mean gain ranged from 21 points in mathematics to 12 points in physics. Campbell and his colleagues also found that those individuals who reported preparing for an Advanced Test did show greater gains. Their conclusion was that subject matter preparation probably led to a real knowledge increase, which in turn was reflected in higher repeated Advanced Test scores.

Since the studies cited above investigated gains across very short time spans, conclusions from these studies must, therefore, be somewhat limited. In the study reported here, time between testings will be treated as a major explanatory variable.

Procedure

A GRE history file containing full records for all persons who had registered for a GRE test between October 1974 and October 1978 was searched for records that had more than one set of scores on the Aptitude Test. Thus, all persons who took the Aptitude Test more than once between 1974 and 1978 are included in the study; in addition, the study includes the records of persons who took the test between 1971 and 1974 if they also registered for another GRE test after October 1974.

It was not possible to identify persons with long time gaps (e.g., more than a maximum of 7 years) between tests through the use of the GRE history file. The alternative was hand searches for scores on microfilm records, which was considered too expensive and time-consuming for this study. Therefore, the aspect of the study concerning the relationship of time between tests and test scores does not consider the long time intervals that were initially proposed, such as time intervals that might be characteristic of a person returning to graduate study 15 or 20 years after receiving the bachelor's degree. However, time intervals of 3 years or more are relatively rare in studies of adult performance on measures of academic aptitude, and it is thought that the results of this study are suggestive of those that would be obtained if samples with longer time intervals were used.

Table 1 gives the number of records that were identified with 2, 3, 4, or 5 sets of Aptitude Test scores on the 1974-78 GRE history tape. Unless indicated otherwise, these are the records that were used in all of the following analyses. Table 1 also reports

Table 1

Mean and Total Gains for Verbal and Quantitative Scores by Administration Populations

No. Aptitude Tests Taken	No. Cases	Score Gains: Mean Gain per Administration		Score Gains: Mean Cumulative Gains	
		V	Q	V	Q
2 Admin.	53,668	27.13	23.14	27.13	23.14
3 Admin.	5,074	18.12	15.36	36.23	30.71
4 Admin.	914	13.20	12.59	39.59	37.77
5 Admin.	217	8.31	10.53	33.23	42.12

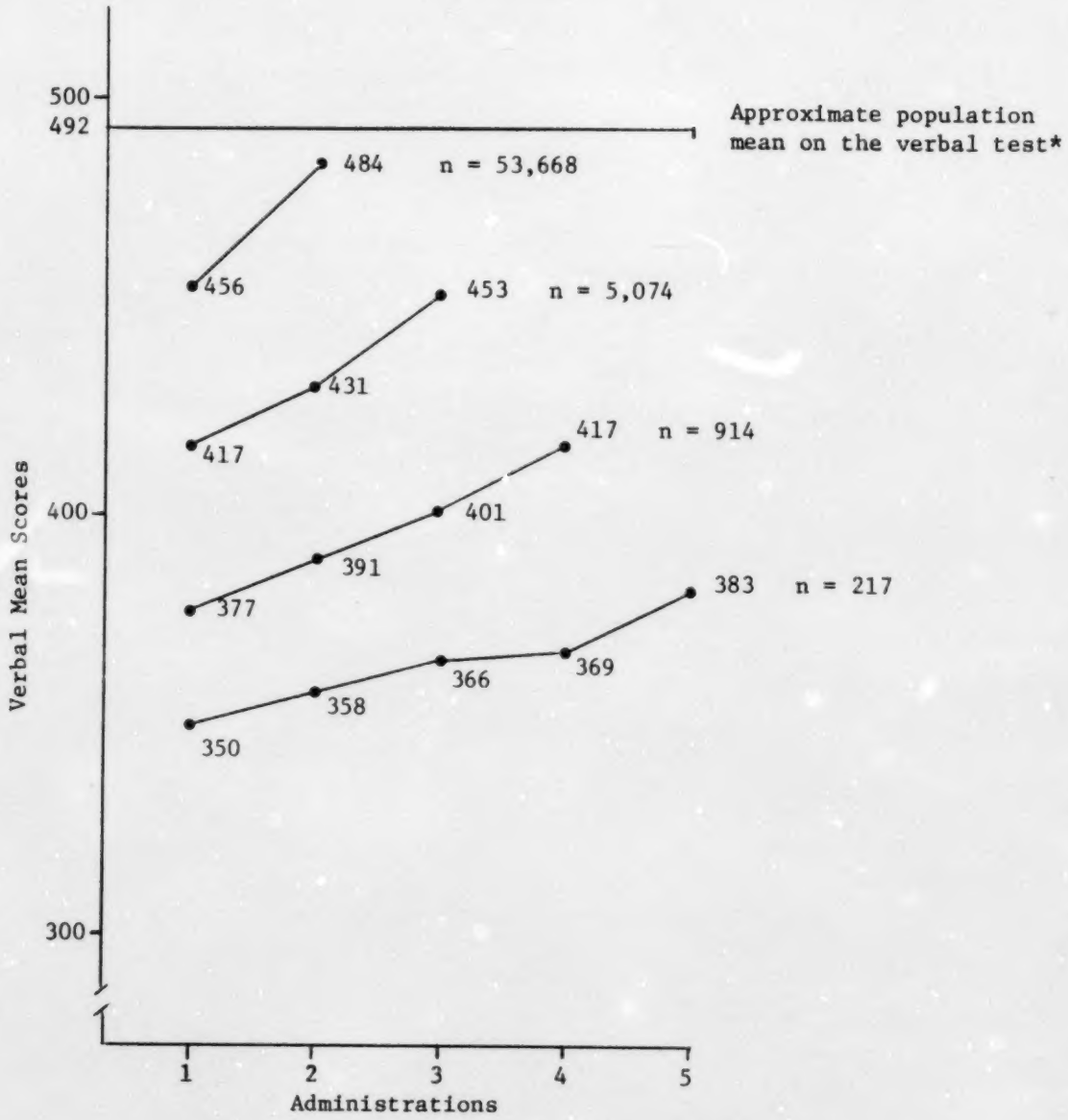
the average score gain on the Verbal and Quantitative Aptitude Tests by administration population.

Descriptive analyses were carried out in an effort to examine changes in the mean levels of scores as well as their rank order changes over time. The basic hypothesis was that long time intervals (more than six months or so) would minimize practice effects, and that the test performance of college graduates would decline with the number of years away from an academic setting. Therefore, it was expected that the greater the time lapse between test and retest, the less would be the mean gain. Similarly, the greater the time lapse, the greater the likelihood that there would be rank order changes, and thus the correlation between test and retest would decrease. For those individuals having three or more retests, the correlations between adjacent (in time) testings should be highest, and the correlations should decrease as the testings become farther apart. When such a systematic pattern is observed among correlations based on repeated measures, we say that the correlations follow a simplex model or pattern (Werts, Linn, & Jöreskog, 1978).

Most simply, test-retest data that follow the simplex pattern suggest that what you knew yesterday is a better estimate of what you know today than what you knew a year ago. Common sense suggests that most developmental models (of which knowledge acquisition is an example) should display a simplex pattern unless (a) the repeated measurements have different reliabilities and thus differentially attenuate the observed correlations; or (b) there is a large gap in time during which previous knowledge is not exercised but later on, because the candidates returned to school or had knowledge-relevant work experience, the previous knowledge is "refreshed," resulting in higher test scores. In such a case, adjacent test scores may not be as highly related to one another as are early test scores that were gathered during college and later test scores that reflect a return to school or relevant work. For example, it is possible that women who retested 4 or more times and who spent considerable time as homemakers and then returned to school may be less likely to show the simplex pattern than men who did not have gaps in their careers.

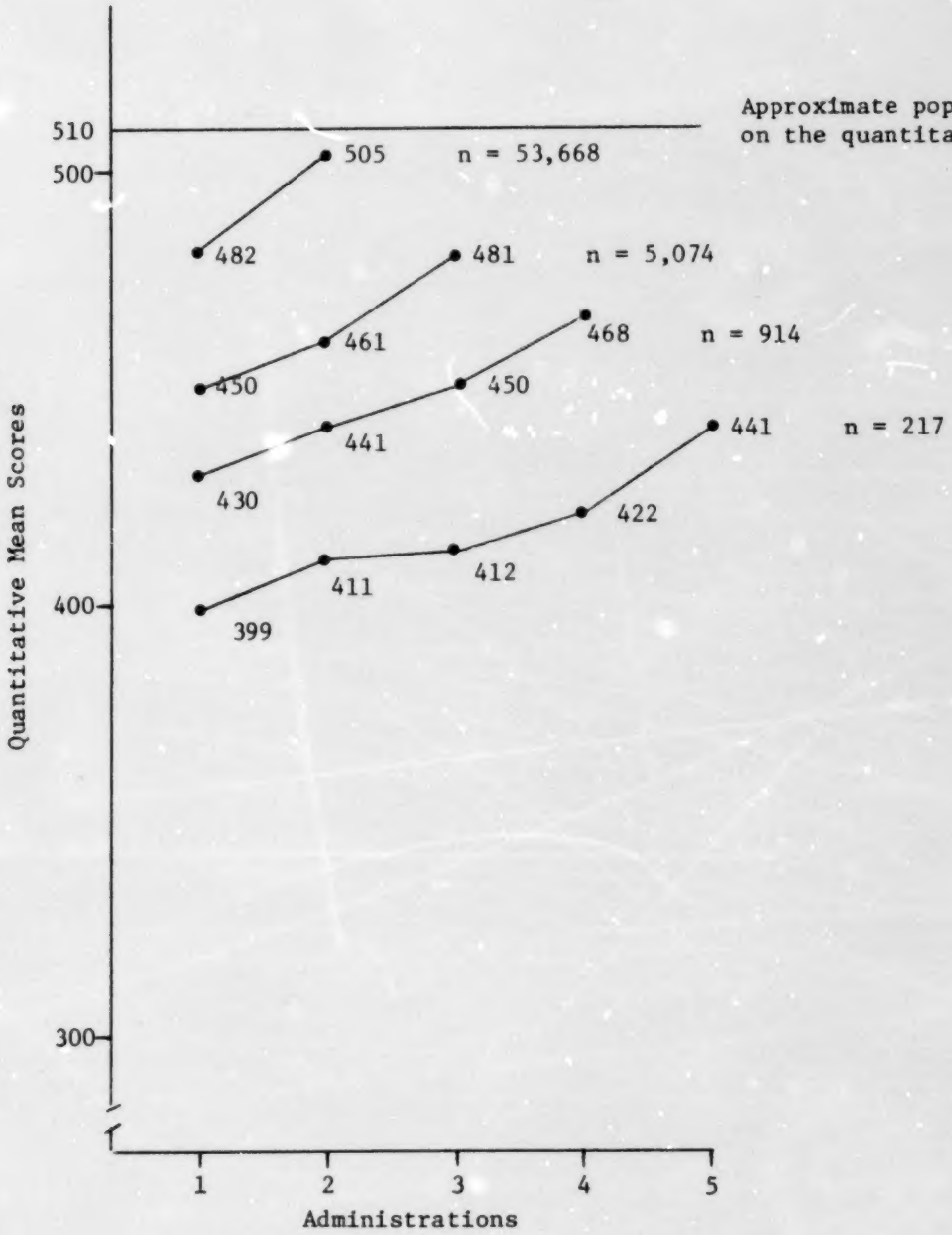
Results and Discussion

Figure 1 presents plots of the repeated verbal mean test scores for each of four populations--those who repeated one, two, three, or four times. Inspection of the verbal score plots indicates that the four populations are quite different both with respect to their initial status and to their final status. The profiles for quantitative scores, presented in Figure 2, show similar patterns. Whether one is a member of the 2-administration population or the 3-, 4-, or 5-administration populations, one can expect to gain on retesting.



*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in test interpretation materials issued by the GRE program.

Figure 1. Verbal mean score changes by number of repeated administrations.



*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in test interpretation materials issued by the GRE program.

Figure 2. Quantitative mean score changes by number of repeated administrations.

However, the average gain per administration in both the verbal and quantitative areas decreases with the number of administrations (see Table 1 for mean gains). With respect to the verbal area, the total or cumulative gain is reasonably close across all administration populations; it just takes the 4- and 5-administration people much more time and practice to arrive at similar total gain. This is especially interesting since the initial status scores for the 4- and 5-administration people average about a standard deviation below the 2-administration people (see Figures 1 and 2). Classical reliability theory would suggest that the 4- and 5-administration people should show greater gains on retesting since they are so far below the population mean. This expectation is based on the fact that some part of the discrepancy between an individual's score and the mean of the population of which he or she is assumed to be a member is attributed to the unreliability of the instrument and to chance performance factors that do not replicate from one administration to another. This is, of course, an over-simplification of the regression phenomenon. In addition, people who score below their group mean become more motivated to achieve at least equal to their peers. Thus one could pose a group social pressure hypothesis to partially explain regression toward group means to supplement the measurement theory. For both of these reasons, we would expect repeaters with the lowest initial test scores to show the largest gains; instead, the persons who repeated the test most often averaged the lowest initial scores and also had the lowest average gain from one administration to the next.

The question arises whether 2-administration individuals who had low initial test scores like those of the 4- and 5-administration individuals gain at the same low rate as the 4- and 5-administration individuals on retesting. The answer is no. For example, the average gain of 2-administration people whose initial test scores are in the 340-380 range is 29.6-32.7 points (see Appendix A). Similar findings apply to the quantitative area (see Appendix B). Clearly the administration subpopulations are different with respect to measured ability. In a sense, the initial scores of the 4- and 5-administration people appear to be closer to their "true" scores than are the initial scores of the 2-administration individuals. It would seem that if a test taker is more than a standard deviation (about 125 points) below the population average at initial testing and does not gain about a quarter of a standard deviation on retest, he or she will probably show only modest gains on each succeeding testing.

What is the expected or mean gain on first retesting combining all administration populations? The expected verbal gain is 26.49, and the expected quantitative gain is 23.13. These are more accurate estimates of what might be expected with a single retesting than are the mean gains reported in Table 1 for the 2-administration group, as they minimize the bias resulting from self-selection

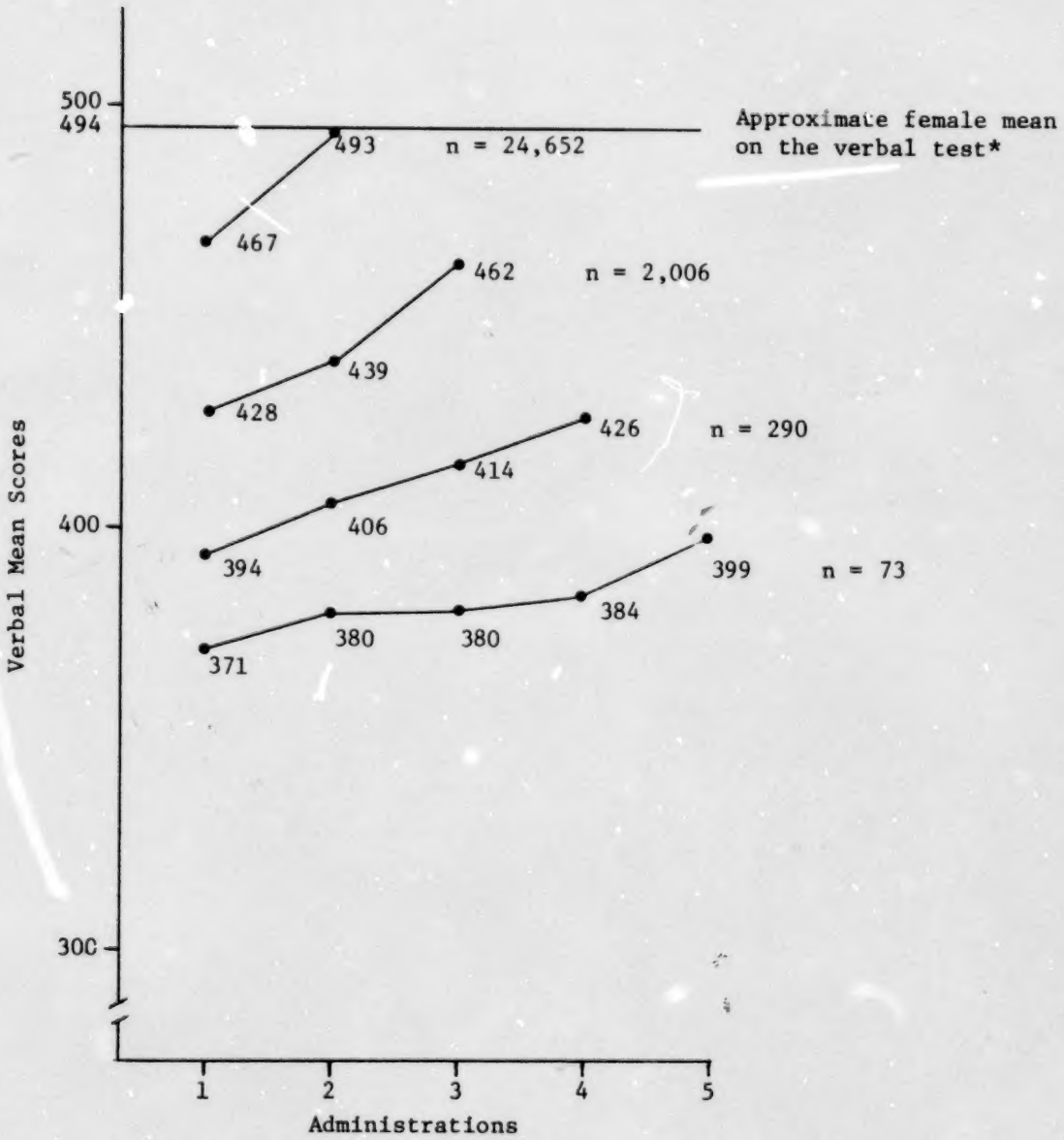
among the administration populations. The reader should note, however, that the bias is extremely trivial, since the number of 2-administration people is almost 9 times larger than the number who repeated the test more than once.

Whether the gains are primarily the result of practice effects or of small gains in knowledge acquisition cannot be conclusively resolved by the data. However, farther on we shall look at the relationship between time lapse and gain, which suggests that at least part of the gain is knowledge acquisition. These initial findings with respect to gain are consistent with the results reported by Campbell, Hilton, and Pitcher (1967). They found gains of 25 and 16 points for the verbal and quantitative areas respectively when retesting was done within three months. Appendices C through F show similar verbal gains (23-24 points) and slightly higher quantitative gains (24-25 points) after a similar time lapse.

Score Gains by Men and Women

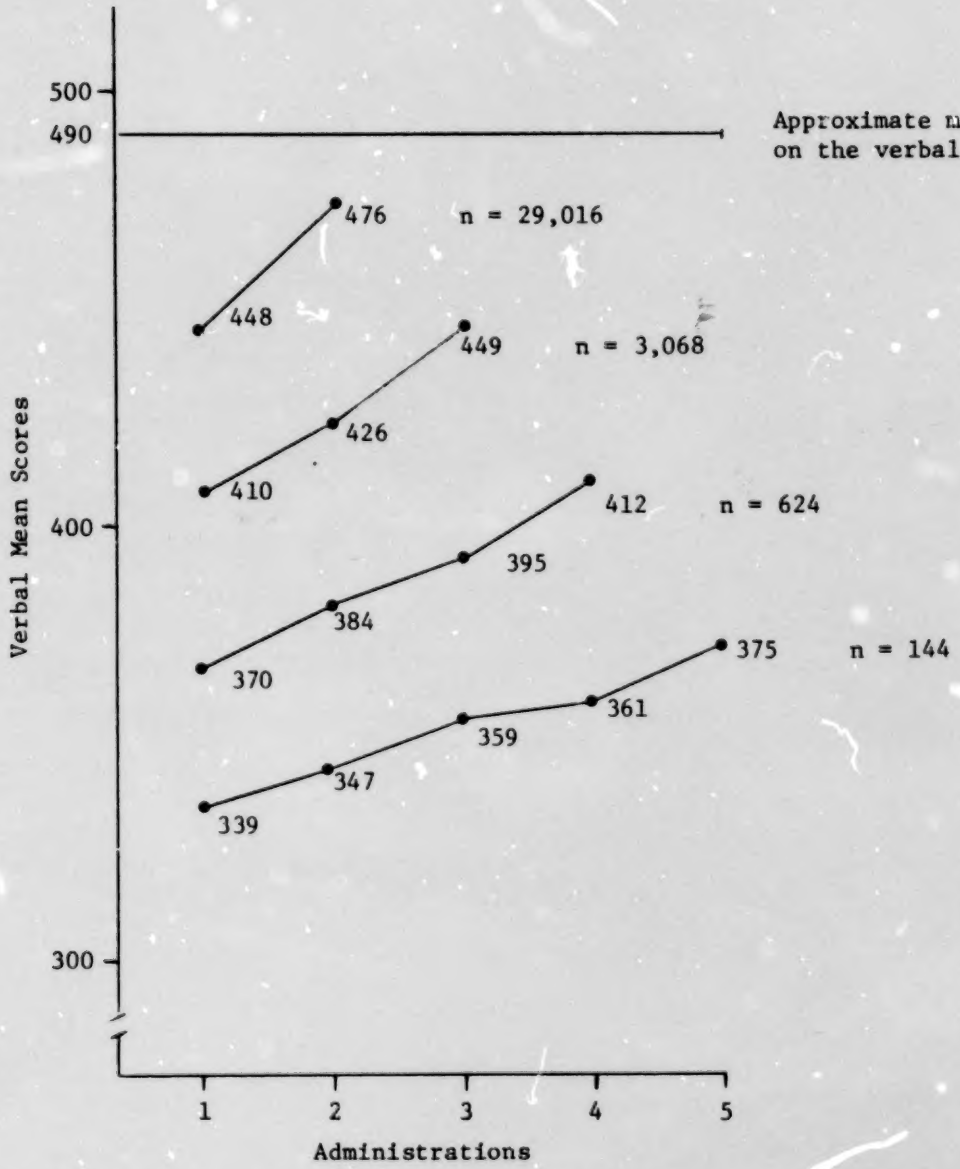
Figures 3 through 6 present plots of initial and retest mean scores by sex and ability area. The profiles appear to be markedly the same for the two sexes. It is interesting to note that proportionately almost as many women as men retested once. However, the number of women becomes proportionately smaller in the multiple-retesting populations.

Table 2 summarizes the gains reported in Figures 3 - 6 by presenting the mean gain between administrations and the cumulative gain for each aptitude area by sex and administration population. For both sexes the average gain between retests decreases with the number of administrations. However, one must avoid over-interpreting the results of the 5-administration individuals since the population size is so small (73 women and 144 men). A comparison of the average gain for the two sexes suggests that men repeaters gain slightly more than women repeaters on both the verbal and quantitative examinations. If we assume that females have as much opportunity to develop their verbal skills between testings, then the 1- to 3-point difference in average verbal score gain may simply be due to regression effects. That is, men repeaters average 42 to 151 points (depending on the administration population--see Figure 4) below the verbal mean of the total male population at initial testing, whereas women at initial testing average 27 to 123 points below their population mean (see Figure 3). Thus if the men are regressing on their population mean and similarly the women are regressing on their population mean, then one would expect a slightly greater gain for the men on retesting simply because of regression effects. It is also interesting to note that 2-administration women actually exceed their population mean on retesting in the quantitative area (see Figure 5).



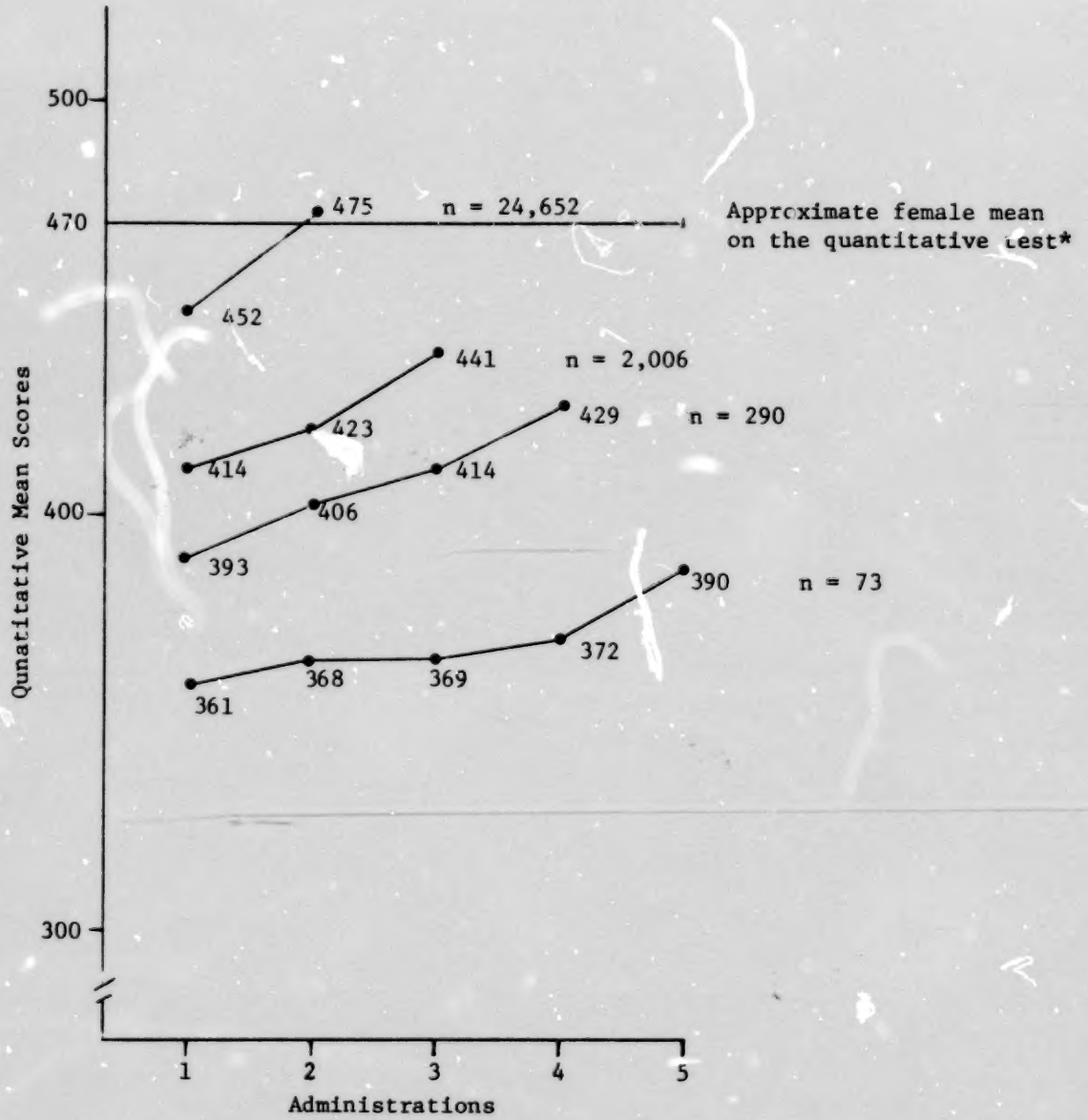
*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in the test interpretation materials issued by the GRE program.

Figure 3. Verbal mean score changes for women by number of repeated administrations.



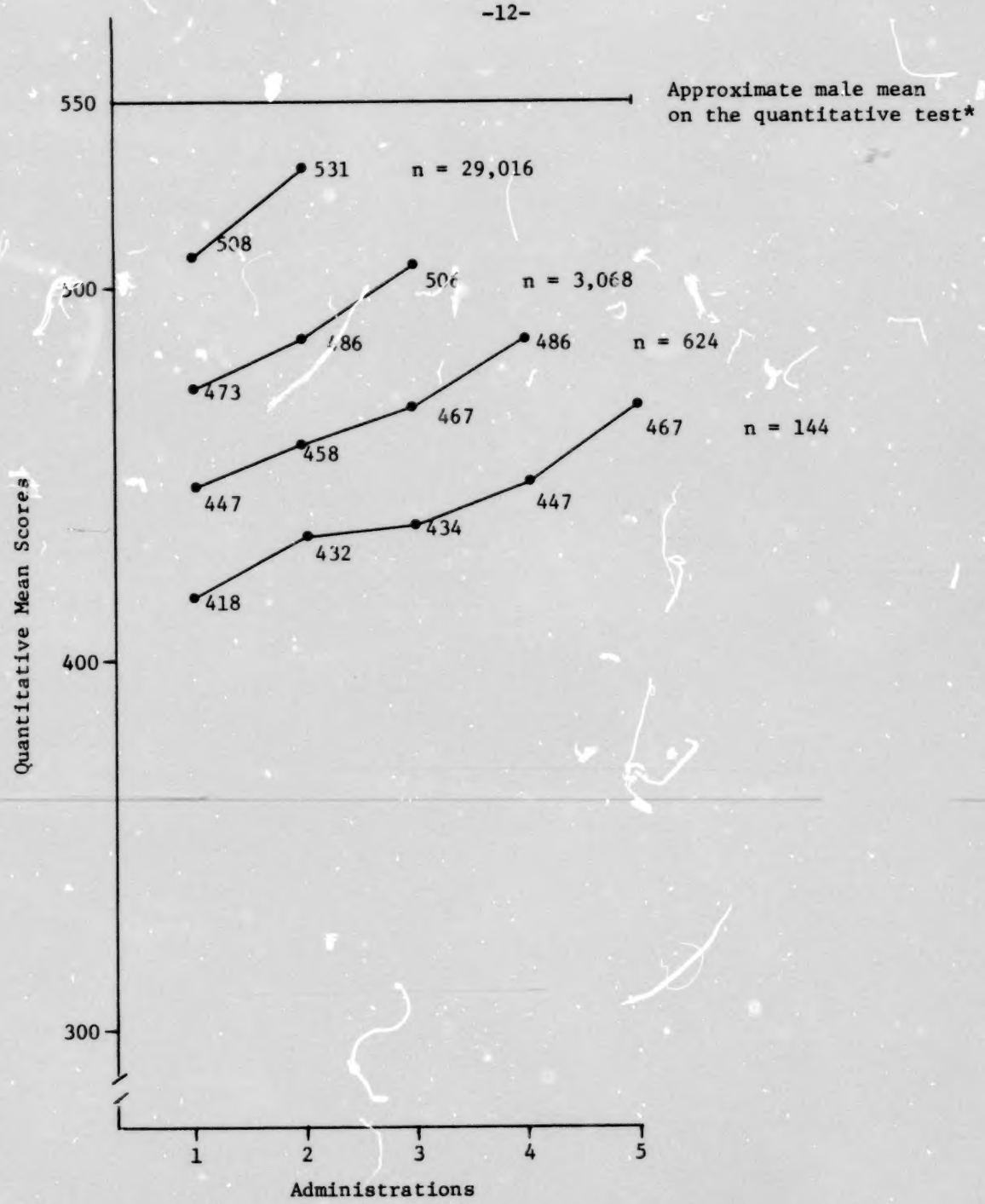
*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in the test interpretation materials issued by the GRE program.

Figure 4. Verbal mean score changes for men by number of repeated administrations.



*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in the test interpretation materials issued by the GRE program.

Figure 5. Quantitative mean score changes for women by number of repeated administrations.



*The approximate population mean is the mean of all test-takers in the early 1970's, as reported in the test interpretation materials issued by the GRE program.

Figure 6. Quantitative mean score changes for men by number of repeated administrations.

Table 2

Mean and Total Gains for Verbal and Quantitative Scores by Sex and Administration Populations

No. admin.	Mean gain between admin.				Cumulative gain			
	Verbal		Quantitative		Verbal		Quantitative	
	M	F	M	F	M	F	M	F
2	27.83	26.13	23.39	22.95	27.83	26.13	23.39	22.95
3	18.99	16.91	16.39	13.37	37.99	33.82	32.79	26.75
4	14.22	10.85	12.88	11.73	42.67	32.55	38.64	35.20
5	9.04	6.85	12.18	7.26	36.18	27.40	48.75	29.04

The fact that men tend to gain slightly more on the average than women on retesting in the quantitative area can also be explained by regression effects since the male administration populations are proportionately farther away from their total population mean than are the female administration populations. It is also possible that more men have an opportunity to maintain and/or develop their quantitative skills between testings. It should be noted, however, that the differential gain by sex, whether verbal or quantitative, is relatively trivial.

Score Gains and Time Between Tests

Both the Levine and Angoff study (1958) and the Weiss review (1961) suggest that gains tend to decrease after either the first or the second testing. For the most part their results were based on short time intervals and thus observed gains were more likely due to practice effects (e.g., improvement in test-taking skills) than to developmental acquisition of knowledge over time. In this section we will show that gains in GRE verbal scores for persons who repeat the test only once also increase as a function of the time span between test and retest. This suggests that the test scores are reflecting the developmental acquisition of verbal knowledge as well as showing gains resulting from possible short-term practice effects.

Table 3 presents the standardized regression weights when retest scores are regressed on both initial test scores and the time lapse between testings. The analysis was done on the 2-administration test takers. For verbal scores there is a small positive effect of

time lapse on gain. This is more prevalent for males than for females. That is, the partial regression weight, holding initial status constant, is positive and non-zero, indicating that larger gains are associated with longer time lapses between testings. Thus, in the verbal area, males are likely to show greater gains between initial and first retest scores as the time lapse increases between testings, regardless of initial test score. There is a similar statistically significant gain in female verbal scores with increasing time intervals, but it is somewhat smaller. These results are contrary to the initial research hypothesis of decreasing scores with long time intervals between tests; instead, they are consistent with a "growth" model of increasing knowledge over time.

Table 3

Regression of Retest Scores on Initial Scores and Time Lapse by Sex and Aptitude Areas

Independent Variables	Dependent Variable			
	Men		Women	
	Standard Score Regression Weights		Standard Score Regression Weights	
	Verbal retest			
Initial Verbal	.8801		.8825	
		R = .885		R = .883
# of months	.0417		.0245	
	Quantitative retest			
Initial Quantitative	.8613		.8482	
		R = .861		R = .848
# of months	-.0110		.0026	

When initial quantitative scores are controlled, time lapse has no relationship to the retest scores, as indicated by the very small and nonsignificant partial regression weights in the lower part of Table 3. This finding is consistent with the idea that gain in mathematics achievement is less likely to occur with the simple passage of time unless one is practicing mathematical skills during that time lapse. These findings are consistent with the study by Owens (1966), who reported that verbal skill increases throughout the adult years, while quantitative skill is relatively stable.

Appendices C through F present mean gains for 2-administration men and women by aptitude area according to time-lapse intervals of 5 months to more than 3 years. This analysis was carried out to (a) detect any nonlinearity in the relationship between time lapse and gain on retesting, and (b) investigate the relative contribution of practice effects to gain. The regression analysis reported in Table 3, of course, assumed linearity. Inspection of Appendices C and D shows a relatively monotonic relationship between time lapse and verbal score gain for both sexes. That is, as the time lapse between testing increases, the gain in verbal test scores also increases. However, in the quantitative area there appears to be no simple relationship between time lapse and retest performance (cf. Appendices E and F). Males in the 2-administration population show the greatest gain in quantitative scores if they retest within five months. This suggests that in the quantitative area practice effects may be a more important determinant of gain for men than is the case in the verbal area. There does not appear to be a similar practice effect for women. This finding is consistent with Weiss' (1966) conclusion that practice effects interact with ability. That is, 2-administration men demonstrate more initial measured mathematics ability and also show greater gain when retested within five months.

Regression Effects and "True" Gain

How much of the 26- to 27-point average gain on the verbal retest and the 23-point average gain on the quantitative test can be ascribed to regression effects? The answer depends on the score level obtained on initial testing. At the extreme lower end of the test-score distribution one would expect a larger component of the gain to be due to regression effects, while at the mean there should not be any regression effects. That is, classical test theory (see Lord & Novick, 1968, p. 65) would suggest that an individual in the 300-310 verbal initial test range should gain about 13 points simply on the basis of regression effects. Inspection of Appendix A indicates that individuals in that range have a mean observed score gain of 32 points. Thus, it would seem that individuals at this tested ability level only show on the average a "true" gain of 19 points--"true" in the sense of a gain beyond that expected

from regression effects. Similarly, an individual with an initial score in the 400-410 range would be expected to gain about 6 points based on regression effects. Appendix A shows a gain of approximately 31 observed score points for such individuals, suggesting that their "true" gain is about 25 score points.

Since the verbal gain for individuals at the population mean (and therefore no regression effects) is on the average 26 points, one might conclude that there is a slight tendency for the better prepared, as indicated by higher initial test scores, to show proportionately greater "true" gains than are shown by the less well prepared individuals. This hypothesis is supported by the fact that observed gains do not drop below 20 points until one gets to the 600 level on initial score, where we would expect such gains to be relatively unlikely since they must overcome regression effects towards a lower mean of approximately 490 (see Appendix A). It would seem that the higher-scoring candidates who decide to retest do so for good reason, and they profit from some of their mistakes on initial testing. This may also be true of the lower scoring candidates, but to a lesser extent.

Since the reliability coefficients of the verbal and quantitative tests are not quite the same (we have assumed a reliability of .93 for the verbal and .91 for the quantitative tests in our calculations here), one would expect the regression effect to be slightly greater for the quantitative test. There is a slighty larger observed average gain for the individuals who scored initially in the 300-400 range on the quantitative test, compared with score gains at this range on the verbal test (see Appendix B). However, because of the reliability difference, the true gain for these initially low-scoring individuals is about the same as for the comparable verbal scores. Conversely, the average gain for individuals above the overall quantitative mean is proportionately somewhat less than that found in the verbal area. In summary then, the "true" gain on retesting for low-scoring individuals in the quantitative area is approximately the same as for similar individuals in the verbal area, or about 20 points. However, the initially higher scoring individuals (590-600 range) on the quantitative tests have "true" gains in the 20-22 point area, whereas individuals scoring in this range in the verbal area gain 29-30 points.

A word of caution is in order concerning these estimates of "true" gain. We have assumed that the reliability of a test is relatively invariant throughout the range of observed scores. One would expect the test to be less reliable at the very low score ranges (because of more guessing) as compared with the 500-600 score range. Thus, our estimates of "true" gain in the 300-400 range may be slightly overestimated. In short, the very low-scoring repeaters may well have a larger error component in their gains than is being estimated above, which would reduce the amount of "true" gain that could be expected.

Characteristics of Repeaters

Appendices G and H present frequency distributions of gains on first retest for verbal and quantitative scores. The distributions of gain are quite similar for each aptitude area, with approximately 70% of the repeaters showing some score gain on retest. These results are consistent with the hypothesis that repeaters are a self-selected group of test takers who have some reason to think that they can improve their scores.

In an effort to identify biographical information that might differentiate single repeaters from multiple repeaters, responses to background information questions were used to predict group memberships (i.e., membership in the single repeating group vs. the multiple repeating group). Appendix I presents the background questionnaire items; asterisks identify the items that were included in the regression analysis. Table 4 presents the results of this stepwise regression.

Table 4

Regression of Repeater Group Membership on Demographic Variables

Independent Variables	Standard Score Regression Weights to Predict Multiple Repeaters	Multiple Correlation
Initial Verbal Score	-.1272	
Degree Aspirations	.0833	R=.166

The stepwise selection procedure was concluded when either the increment in the multiple correlation was less than .01 or the standard score regression weight was less than .08. These somewhat arbitrary decisions were implemented because the large sample size made statistical tests relatively meaningless. Inspection of the standardized weights suggests that, after controlling for measured ability, a multiple repeater is more likely to have a higher degree objective. Clearly there is considerable unexplained variance with respect to who becomes or does not become a multiple repeater, as indicated by a multiple correlation of only .295. Appendix J presents the intercorrelations, means, and standard deviations for the variables used in the analyses reported in Table 4 and Table 5.

The biographical information was also used to predict changes in rank ordering from initial test to first retest. That is, are demographic variables related to whether an individual goes up or down from test to retest? Table 5 presents results of stepwise regressions of verbal and quantitative retest scores on initial scores and "significant" demographic variables.

Table 5

Regressions of Verbal and Quantitative Retest Scores on Initial Scores and Significant Demographic Variables

Independent Variables	Verbal Retest Standard Score Regression weights		Quantitative Retest Standard Score Regression weights	
Initial Verbal Score	.8525		.1252	
Initial Quantitative Score	.0785	R=.8891	.8051	R=.8708
U.S. citizen (yes=1, no=2)			.0880	

The same criteria discussed above were used for inclusion of a variable in the final equation. As in the case of predicting repeater group membership, demographic information adds little to the prediction of rank-order changes on retesting. However, one interesting comparison between gains on the verbal and quantitative areas can be made. For the quantitative area, test takers are more likely to show slightly greater gains when controlling for initial status if they are not U.S. citizens (i.e., for the most part, they use English as a second language). The opposite appears to be true in the verbal area. That is, although the verbal regression for U.S. citizens was not quite "significant" by the stipulated criterion (-.051), it was negative in sign, indicating that non-U.S. citizens gain less than U.S. citizens on retesting in the verbal area. The reader should note that this is probably an underestimate of the effect of citizenship since the partial regression weight tends to underestimate the effect of a dichotomous variable when the split is extreme. In this case approximately 89% of the retesters are U.S. citizens, while only 11% report that they are non-U.S. citizens.

Because of such "slim splits" on some of the demographic variables and the resulting possible underestimate of their relationship to score gain, additional analyses of gains were carried out for these variables. Table 6 presents mean gains for selected subgroups whose effects were likely to be underestimated in the regression analysis.[1]

Inspection of Table 6 indicates that there is a slight difference in favor of the U.S. citizens with respect to average gain on the verbal test. However, on the quantitative test the non-U.S. citizens average a gain 11 points higher than the gain of U.S. citizens. This result is, of course, consistent with results of the stepwise multiple regression analysis reported in Table 5. It is possible that there may be a quantitative "practice" effect for non-U.S. citizens in the sense that they may not be as familiar as U.S. citizens with the objective test format and item types. It is also likely that a slight improvement in language facility may be just enough to bring a foreign student's language capabilities to the minimum level necessary to perform well on the quantitative test. This gain is even more impressive when one considers that non-U.S. citizens' initial quantitative test scores average 22 points higher than the initial mean scores of U.S. citizens. Clearly, the non-U.S. citizen is an example of the type of individual who because of language limitations may well expect to gain on repeating the quantitative test.

Table 6 also shows that the Black mean score gain on first retest in the verbal area is approximately 10 points less than that of the White population and 6 points less than that of "Other" groups. One might expect that part of this disparity reflects the fact that the initial verbal mean score of Black repeaters is closer to the overall Black verbal mean score than is the case for White repeaters. That is, the Black repeaters' initial verbal score (365.06) is only about 5 points less than the overall Black mean. Thus, we might expect almost no regression effects for Black repeaters. However, members of the total repeater population with initial verbal scores near the mean verbal score for the total GRE population still gained approximately 26 points on retesting (e.g., Figure 1 and Appendix A). Therefore, the explanation of this differential gain on the basis of regression effects appears unlikely. What seems more likely is that the higher scoring repeaters are more likely to profit from their previous mistakes and thus show proportionately larger "true" gains on retesting--that is, larger gains

---

[1]Since questionnaire data were available only for test takers in 1976-1978, the number of cases available for these analyses was much smaller than for earlier analyses.

Table 6

Mean Scores on Initial Testing and Mean Gains on First Retesting for Selected Subgroups

Groups	N	Verbal			Quantitative		
		Mean Score on Initial Testing	Mean Gain on Retesting	Standard Deviation of the Gains	Mean Score on Initial Testing	Mean Gain on Retesting	Standard Deviation of the Gains
U.S. Citizen	9838	474.71	23.80	56.34	492.33	20.38	66.59
Not U.S. Citizen	1247	341.38	21.33	62.71	514.60	31.30	73.60
White	8077	483.93	24.76	56.50	504.40	20.68	66.45
Black	684	365.06	14.77	55.99	360.96	18.87	67.02
Others	1138	425.81	21.05	58.17	489.48	23.17	69.34

independent of regression artifacts. Another possibility is that many Blacks may be encouraged by peers (or admissions officers) to repeat tests even though their initial test scores are reasonably representative of their present performance level on academic tasks of the types that are covered by the tests.

Simplex Model

Up to this point, we have investigated changes in level or mean gains (with the exception of the regression analysis) when individuals repeat the GRE Aptitude Test. Appendices K, L, and M present the correlational tables that underlie the simplex model and that are concerned with changes in rank order over repeated retestings. Men and women with 4 administrations were combined in Appendix K; Appendices L and M show correlations by aptitude area within sex. Inspection of the correlations suggests that there is little or no change in rank order for the verbal testings. The small observed changes probably result from differences in time lapse among the individuals being retested. That is, one individual might have one year between his second and third testing, while another individual might have only 4 months. The earlier results from the regression analysis (Table 3 and Appendices C and D) suggest that individuals tend to gain more on verbal retesting with longer time lapses between testings.

The quantitative retest scores are only slightly less correlated with one another than are the verbal retest scores. In both the verbal and the quantitative area, when the sexes are combined, the pattern of correlations follows the simplex model with no reversals. That is, correlations adjacent in time are the highest, and as the testings become more distant, the correlations become smaller.

Inspection of Appendices K and L, however, shows reversals in the expected simplex pattern for females. The asterisks indicate three reversals in the quantitative area and one in the verbal area for females. There are no reversals for males. The extent of female reversals is small, and, considering the sample size (n = 290), any conclusions would at best be very tentative. The reversals, however, are consistent with the notion that females may show a deviation from simplex patterns (particularly in mathematics) because of less consistent opportunities to maintain their quantitative skills between retestings.

Summary and Conclusions

The objectives of this study were to obtain information on (a) how many persons repeat the Graduate Record Examinations (GRE) Aptitude Test one or more times; (b) whether there is a time-related decrement or increment for those who repeat, (c) whether the pattern of increment or decrement differs by sex and/or by ability area, and (d) whether single repeaters differ from multiple repeaters in ability and/or sex group membership. In order to answer these questions, descriptive analyses were carried out on all repeaters whose most recent administration was between October 1974 and September 1978. The analyses were of two types. The first type of analysis was concerned with changes in mean scores over testings by sex group and aptitude area. The second type of analysis was concerned with an individual's change in rank order over testings.

It was found that, on the average, individuals who repeated the GRE Aptitude Test improved their scores on retesting, regardless of the aptitude area or sex group. Individuals who retested only once (the vast majority) could be expected to gain an average of about one-fourth of a standard deviation on the verbal test (about 26-27 points) and only slightly less (about 23 points) on the quantitative test. Individuals who retested more than once scored lower on the average on initial testing than did the single retest individuals. Furthermore, the multiple retesters gained less on the average from test to test than did the higher scoring single retest individuals. This finding is, of course, consistent with the expectation that if a person does not gain significantly on the first retest, he or she is more likely to continue to retest. In addition, a close examination of "true" gains on first retest indicates that individuals with higher initial scores (i.e., in the range of 500-600) have proportionately greater "true" gain than those who are in the 300-400 range on initial testing, where "true" gain is defined as gains corrected for regression effects that result from the random errors of testing. Thus, it appears that higher scorers are better able to estimate the likelihood of improving their scores when deciding whether or not to repeat the test. Clearly, repeaters are a self-selected group who have some reason to believe that the first testing did not accurately reflect their ability. Because of these self-selection factors, they are more likely to obtain a higher score on retesting than would be the case for a random sample of all test takers.

The total or cumulative gain of the lower scoring individuals in the multiple retest population was equal to or slightly greater than the mean gain of the single retest population. Apparently these multiple retest people take a longer time and more practice to achieve total gains similar to those of the single retest population. However, the final mean scores of multiple repeaters remained significantly lower than those of the single repeaters; they appear to be from a population of lower ability and may be regressing on a different mean.

Both males and females demonstrated, with certain minor exceptions, the same overall patterns of gain. The average gains over retestings were slightly in favor of the males. This slight difference could be explained by regression toward different group means. That is, the initial test scores of male retesters were farther below the male population mean scores than were the corresponding scores for the females and thus one would expect greater gains for the males on the basis of regression effects. Women were proportionately less persistent than men with respect to continuing to retest. Both men and women single repeaters showed greater verbal gains when there were longer time lapses (up to 3 years or more) between testings, contrary to our initial hypothesis of some score loss with long time spans between testings. It was concluded that much of these retest gains in the verbal area reflected increased knowledge and not just practice effects. There was no apparent relationship between time lapse and quantitative gains. For men only, there was some indication that the amount of expected increase might decline with longer time intervals (e.g., more than 2 years) between tests, but there continued to be average gains at every time interval for both sexes.

Gains that could reasonably be attributed only to short-term practice effects appeared on the quantitative test, and only for men. That is, men in the single retest population showed the greatest gain in quantitative scores if retested within five months or less. This was not true for women.

Analysis of non-U.S. citizens showed that they made greater gains on retesting in the quantitative area than did the U.S. citizens. There was little difference between U.S. and non-U.S. citizens with respect to verbal gains. Black repeaters showed less verbal gain on the average than did the White and "Other" repeaters. All racial groups showed about the same average retest gain on the quantitative test.

Tables were compiled that report the mean gain of single repeaters at various initial test score levels (see Appendices A and B). These may be of use to test takers in deciding whether or not to repeat the test, and to admissions directors in interpreting multiple test scores.

One should not conclude from the above findings that every GRE candidate if retested could expect to gain 20 to 27 points. Individuals who retest are self-selected in many unknown ways. Many may decide to retest because they did not finish the first time they took the test, because of poor time scheduling, because they feel they made mistakes that they could rectify on re-examination, or because they feel they know more. In other words, they elect to take the test again because they have some reason to think that they can improve their scores. The results reported in this paper confirm their expectations.

References

Campbell, J. T., Hilton, T. L., & Pitcher, B. Effects of repeating on test scores of the graduate record examinations (GRE Special Report 67-1). Princeton, N.J.: Educational Testing Service, April 1967.

Levine, R. S., & Angoff, W. H. The effects of practice and growth on scores on the Scholastic Aptitude Test (ETS SR-58-6). Princeton, N.J.: Educational Testing Service, February 1958.

Lord, F. M., & Novick, M. Statistical theories of mental tests scores. Reading, Mass.: Addison-Wesley, 1968.

Owens, W. A. Age and mental abilities: A second adult follow-up. Journal of Educational Psychology, 1966, 57, 311-325.

Weiss, R. A. The effects of practicing a test: A review of the literature (ETS RM-61-12). Princeton, N.J.: Educational Testing Service, November 1961.

Werts, C. E., Linn, R. L., & Jöreskog, K. G. A simplex model for analyzing academic growth. Educational and Psychological Measurement, 1977, 37, pp. 745-756.

Appendix A

Verbal Score Changes on Retesting by Level of Initial Score
(Single Repeaters Only)

Initial GRE-V Score	N	Mean Gain	SD(N-1)	Range of Score Changes	
				Low	High
200	365	62.2	51.0	0.0	330.0
210	201	60.3	67.3	-10.0	530.0
220	330	52.0	60.1	-20.0	540.0
230	311	47.8	57.8	-30.0	370.0
240	521	40.9	59.6	-40.0	440.0
250	425	35.1	52.9	-50.0	280.0
260	729	33.8	59.4	-60.0	470.0
270	559	27.0	54.8	-70.0	230.0
280	700	27.0	54.5	-80.0	240.0
290	808	28.6	62.1	-90.0	410.0
300	628	32.4	57.9	-100.0	340.0
310	941	33.1	56.5	-110.0	310.0
320	837	30.4	57.6	-120.0	390.0
330	1110	29.4	57.1	-130.0	330.0
340	857	29.6	58.8	-130.0	220.0
350	1295	30.3	57.9	-150.0	370.0
360	1221	32.4	58.5	-160.0	280.0
370	1206	29.7	56.4	-160.0	390.0
380	1636	32.7	58.7	-180.0	360.0
390	1352	30.1	58.2	-180.0	320.0
400	1835	30.7	57.3	-200.0	220.0
410	1389	30.6	57.2	-150.0	270.0
420	1681	31.4	56.8	-140.0	250.0
430	1704	30.4	59.0	-170.0	260.0
440	1802	27.0	57.4	-180.0	340.0
450	1988	31.2	56.5	-160.0	250.0
460	1777	27.0	55.4	-200.0	270.0
470	2067	29.5	57.6	-230.0	330.0
480	1584	27.9	58.4	-170.0	260.0

Appendix A (Continued)

Initial GRE-V Score	N	Mean Gain	SD(N-1)	Range of Score Changes Low	High
490	2084	26.3	55.4	-150.0	260.0
500	1504	26.8	57.0	-190.0	230.0
510	1742	26.2	58.3	-310.0	230.0
520	1385	27.9	57.3	-170.0	240.0
530	1645	24.2	56.3	-170.0	230.0
540	1480	24.3	56.1	-170.0	200.0
550	1213	22.6	56.8	-160.0	250.0
560	1131	23.2	57.8	-250.0	240.0
570	1293	21.7	56.9	-210.0	200.0
580	1080	21.1	56.8	-330.0	230.0
590	1031	22.4	53.8	-370.0	190.0
600	911	17.8	56.6	-200.0	180.0
610	664	14.9	54.0	-200.0	190.0
620	667	21.6	51.7	-200.0	160.0
630	626	13.0	52.2	-170.0	160.0
640	598	15.7	51.7	-150.0	170.0
650	435	17.4	52.5	-170.0	150.0
660	478	8.7	49.9	-190.0	160.0
670	350	15.5	49.8	-140.0	140.0
680	282	8.0	50.1	-140.0	140.0
690	320	6.6	47.9	-140.0	140.0
700	149	7.7	48.1	-120.0	100.0
710	196	- 1.3	48.0	-140.0	100.0
720	114	8.3	42.8	-100.0	100.0
730	106	3.3	42.7	-110.0	110.0
740	84	- 5.0	45.0	-100.0	80.0
750	59	-15.4	40.7	-110.0	70.0
760	50	-19.8	88.8	-560.0	70.0
770	42	- 3.6	38.8	-110.0	50.0

Appendix B

Quantitative Score Changes on Retesting by Level of Initial Score
(Single Repeaters Only)

Initial GRE-Q Score	N	Mean Gain	SD(N-1)	Range of Score Changes Low	High
200	343	58.3	61.9	0.0	390.0
210	213	54.0	56.7	-10.0	380.0
220	293	52.6	59.2	-20.0	400.0
230	208	53.2	67.9	-30.0	450.0
240	344	46.0	69.7	-40.0	420.0
250	377	47.0	62.2	-50.0	300.0
260	418	52.5	63.9	-60.0	330.0
270	277	38.1	63.2	-70.0	250.0
280	479	42.6	65.4	-80.0	310.0
290	565	37.5	63.8	-90.0	380.0
300	707	39.8	67.9	-100.0	330.0
310	749	36.2	66.1	-110.0	450.0
320	895	36.4	65.9	-120.0	280.0
330	799	38.2	64.5	-130.0	260.0
340	860	37.4	69.5	-140.0	330.0
350	707	37.8	65.6	-150.0	320.0
360	1120	32.8	66.8	-160.0	340.0
370	1224	34.3	66.0	-170.0	320.0
380	1338	35.0	66.6	-180.0	270.0
390	1492	34.9	66.0	-190.0	300.0
400	1324	31.3	65.7	-160.0	330.0
410	1281	33.7	65.6	-210.0	370.0
420	1371	30.5	65.4	-180.0	270.0
430	1750	26.3	61.9	-180.0	270.0
440	1659	28.0	65.1	-180.0	270.0
450	1657	27.3	63.4	-160.0	240.0
460	1962	28.5	64.6	-200.0	280.0
470	1497	28.2	64.8	-190.0	300.0
480	1624	22.0	64.1	-180.0	250.0

Appendix B (Continued)

Initial GRE-Q Score	N	Mean Gain	SD(N-1)	Range of Score Changes Low	High
490	1434	24.6	64.0	-180.0	260.0
500	1808	19.4	63.0	-180.0	240.0
510	1596	18.8	62.7	-310.0	280.0
520	1706	18.4	65.9	-240.0	220.0
530	1880	21.0	65.1	-290.0	250.0
540	1205	18.0	62.5	-200.0	220.0
550	1481	18.2	63.9	-220.0	230.0
560	1524	15.7	65.6	-360.0	220.0
570	1220	13.5	65.3	-290.0	200.0
580	1460	10.9	63.8	-210.0	220.0
590	1227	14.5	62.9	-340.0	210.0
600	717	4.1	60.9	-190.0	200.0
610	1020	7.1	61.2	-180.0	200.0
620	1040	6.5	62.4	-300.0	190.0
630	918	6.9	62.5	-280.0	190.0
640	753	8.3	58.4	-170.0	190.0
650	700	4.1	59.3	-170.0	180.0
660	679	10.9	59.8	-210.0	180.0
670	565	- 0.7	64.4	-470.0	160.0
680	597	2.2	57.1	-190.0	150.0
690	479	1.8	59.7	-490.0	140.0
700	326	4.0	59.6	-280.0	150.0
710	400	4.4	51.7	-190.0	140.0
720	243	- 8.2	54.1	-190.0	110.0
730	268	-12.0	57.4	-290.0	90.0
740	208	- 4.2	50.4	-180.0	90.0
750	186	- 3.6	48.9	-160.0	110.0
760	196	- 9.8	55.4	-230.0	90.0
770	128	-10.3	55.3	-440.0	80.0

Appendix C

Verbal Score Gains for Males by Time Lapse

(Single Repeaters Only)

Number of months between test and retest	N	Mean Gain	SD(N-1)
1 - 5	10,122	23.8	56.5
6 - 10	6,551	26.6	57.4
11 - 15	5,066	28.7	56.0
16 - 20	1,941	30.1	60.0
21 - 25	2,912	35.4	59.1
26 - 30	1,372	37.1	57.9
31 - 35	494	37.2	68.8
36 +	538	36.2	63.3

Appendix D

Verbal Score Gains for Females by Time Lapse

(Single Repeaters Only)

Number of months between test and retest	N	Mean Gain	SD(N-1)
1 - 5	9,558	23.7	55.9
6 - 10	5,736	25.3	56.1
11 - 15	4,214	28.1	56.0
16 - 20	1,454	28.9	60.2
21 - 25	1,961	30.2	58.6
26 - 30	1,010	32.1	57.1
31 - 35	327	30.5	59.7
36 +	402	27.3	56.0

Appendix E

Quantitative Score Gains for Males by Time Lapse
(Single Repeaters Only)

Number of months between test and retest	N	Mean Gain	SD(N-1)
1 - 5	10,122	25.1	65.8
6 - 10	6,551	22.5	64.5
11 - 15	5,066	23.0	65.8
16 - 20	1,941	24.1	65.8
21 - 25	2,912	22.3	66.9
26 - 30	1,372	20.5	62.6
31 - 35	494	18.9	74.7
36 +	538	16.0	67.3

Appendix F

Quantitative Score Gains for Females by Time Lapse

(Single Repeaters Only)

Number of months between test and retest	N	Mean Gain	SD(N-1)
1 - 5	9,558	23.4	64.5
6 - 10	5,736	21.4	63.7
11 - 15	4,214	23.3	65.0
16 - 20	1,454	23.9	63.8
21 - 25	1,961	23.4	66.6
26 - 30	1,010	22.7	66.4
31 - 35	327	21.3	67.4
36 +	375	28.5	66.2

GRE REPEATER DATA

Appendix G

Frequency Distribution of Verbal Gains on First Retesting

	TOTAL			
SCORE INTERVALS	F	PCT	CF	C-PCT
MORE THAN 200	171	0.3	59857	100.0
200.000	57	0.1	59686	99.7
190.000	82	0.1	59629	99.6
180.000	117	0.2	59547	99.5
170.000	169	0.3	59430	99.3
160.000	230	0.4	59261	99.0
150.000	380	0.6	59031	98.6
140.000	516	0.9	58651	98.0
130.000	751	1.3	58135	97.1
120.000	1003	1.7	57384	95.9
110.000	1269	2.1	56381	94.2
100.000	1621	2.7	55112	92.1
90.000	2101	3.5	53491	89.4
80.000	2463	4.1	51390	85.9
70.000	3093	5.2	48927	81.7
60.000	3552	5.9	45834	76.6
50.000	3817	6.4	42282	70.6
40.000	4163	7.0	38465	64.3
30.000	4253	7.1	34302	57.3
20.000	4403	7.4	30049	50.2
10.000	4052	6.8	25646	42.8
0.0	4046	6.8	21594	36.1
-10.000	3630	6.1	17548	29.3
-20.000	3131	5.2	13918	23.3
-30.000	2667	4.5	10787	18.0
-40.000	2125	3.6	8120	13.6
-50.000	1714	2.9	5995	10.0
-60.000	1360	2.3	4281	7.2
-70.000	913	1.5	2921	4.9
-80.000	673	1.1	2008	3.4
-90.000	488	0.8	1335	2.2
-100.000	295	0.5	847	1.4
-110.000	206	0.3	552	0.9
-120.000	112	0.2	346	0.6
-130.000	87	0.1	234	0.4
-140.000	55	0.1	147	0.2
-150.000	26	0.0	92	0.2
-160.000	17	0.0	66	0.1
-170.000	21	0.0	49	0.1
-180.000	10	0.0	28	0.0
-190.000	2	0.0	18	0.0
-200.000	7	0.0	16	0.0
LESS THAN 200	9	0.0	9	0.0

GRE REPEATER DATA

Appendix H

Frequency Distribution of Quantitative Gains on First Retesting

	TOTAL			
SCORE INTERVALS	F	PCT	CF	C-PCT
MORE THAN 200	228	0.4	59857	100.0
200.000	107	0.2	59629	99.6
190.000	134	0.2	59522	99.4
180.000	192	0.3	59388	99.2
170.000	268	0.4	59196	98.9
160.000	402	0.7	58928	98.4
150.000	473	0.8	58526	97.8
140.000	666	1.1	58053	97.0
130.000	882	1.5	57387	95.9
120.000	1091	1.8	56505	94.4
110.000	1336	2.2	55414	92.6
100.000	1656	2.8	54078	90.3
90.000	2050	3.4	52422	87.6
80.000	2388	4.0	50372	84.2
70.000	2820	4.7	47984	80.2
60.000	3064	5.1	45164	75.5
50.000	3358	5.6	42100	70.3
40.000	3530	5.9	38742	64.7
30.000	3733	6.2	35212	58.8
20.000	3751	6.3	31479	52.6
10.000	3825	6.4	27728	46.3
0.0	3671	6.1	23903	39.9
-10.000	3417	5.7	20232	33.8
-20.000	3081	5.1	16815	28.1
-30.000	2770	4.6	13734	22.9
-40.000	2256	3.8	10964	18.3
-50.000	1965	3.3	8708	14.5
-60.000	1660	2.8	6743	11.3
-70.000	1308	2.2	5083	8.5
-80.000	999	1.7	3775	6.3
-90.000	822	1.4	2776	4.6
-100.000	592	1.0	1954	3.3
-110.000	414	0.7	1362	2.3
-120.000	288	0.5	948	1.6
-130.000	213	0.4	660	1.1
-140.000	147	0.2	447	0.7
-150.000	87	0.1	300	0.5
-160.000	69	0.1	213	0.4
-170.000	46	0.1	144	0.2
-180.000	37	0.1	98	0.2
-190.000	24	0.0	61	0.1
-200.000	7	0.0	37	0.1
LESS THAN 200	30	0.1	30	0.1

Appendix I

Background Information Questions

Your answers to these questions will be used for research and, if you are taking the GRE, in group statistics describing GRE student populations. In both of these uses, individual responses will not be communicated to any institution. In addition, your responses will not affect your scores in any way. If you are registering for the Minority Graduate Student Locater Service, your individual responses to several of the questions may be reported to one or more institutions.

If you are completing the registration form to sign up for the tests only, we encourage you to answer all questions, particularly questions A through N. (You may, however, omit any you do not wish to answer.) *If you are registering for the Locater Service, however, it is essential that you respond to all 20 questions (A through T).*

A. Have you previously taken one or more GRE tests?

(1) No
(2) Yes—took the test(s) on or prior to September 30, 1976
(3) Yes—took the test(s) more recently than September 30, 1976

***If your answer to question A is (3), and if your responses to the rest of the questions would be the same as they were before, you need not answer the questions again* unless you are registering for the Locater Service.** *If your responses to any of the rest of the questions would be different, please respond again to all of them.*

*****B. *Are you a citizen of the United States?*

(1) Yes (2) No

Omit questions C and D if you are not a United States citizen.

C. Referring to the State Code List on page 4, find the code number for the state you consider your permanent residence. Then, fill in the appropriate spaces.

*****D. How do you describe yourself?

(1) American Indian, Eskimo, or Aleut
(2) Black or Afro-American or Negro
(3) Mexican-American or Chicano
(4) Oriental or Asian-American
(5) Puerto Rican
(6) Other Hispanic or Latin-American
(7) White or Caucasian
(8) Other

*****E. Do you communicate better in English than in any other language?

(1) Yes (2) No

*****F. Approximately how many full-time undergraduate students attend the school from which you received or will receive your bachelor's degree?

(1) Fewer than 1,000
(2) 1,000-4,999
(3) 5,000-9,999
(4) 10,000-19,999
(5) 20,000 or more

*****G. Which of the following best describes your undergraduate institution?

(1) Public
(2) Private—no church affiliation
(3) Private—church affiliation

H. In what year did you receive or do you expect to receive your bachelor's degree?

(Please blacken the spaces on your registration form corresponding to the last two digits of the year.)

I. Referring to the Major Field Code List on page 3, find your undergraduate major field of study. Blacken the spaces for that field's code number. If your major is not listed, select one of the following codes and blacken the corresponding spaces:

10 Other Foreign Languages
98 Other Humanities
80 Other Social Sciences
30 Other Biological Sciences
60 Other Physical Sciences
02 Not included in above categories

*****J. What is your eventual graduate degree objective?

(1) Nondegree study
(2) Master's (M.A., M.S., M.Ed., etc.)
(3) Intermediate (such as Specialist)
(4) Doctorate (Ph.D., Ed.D., etc.)
(5) Postdoctoral study

K. Referring to the Major Field Code List on page 3, find the field in which you plan to do your graduate work. Blacken the spaces for that field's code number. If your intended major is not listed, or if you are undecided, select one of the following codes and blacken the corresponding spaces:

10 Other Foreign Languages
98 Other Humanities
80 Other Social Sciences
30 Other Biological Sciences
60 Other Physical Sciences
02 Not included in above categories
00 Undecided

L. If you have a second choice of graduate major field, enter its two-digit code number in the appropriate spaces, following the instructions for question K. If you have no second choice, leave this question blank.

M. Which of the following best describes the graduate institution you most recently attended or currently attend on at least a half-time basis?

(1) I have never attended graduate school or have attended graduate school on less than a half-time basis only. (Skip to O.)
(2) Public
(3) Private—no church affiliation
(4) Private—church affiliation

N. In what year did you last attend graduate school on at least a half-time basis?

(Blacken the spaces on your registration form corresponding to the last two digits of the year, if you have not attended

Appendix I (Continued)

*O. In courses in your undergraduate major field only, what grade average have you received so far? (If your college does not use letter grades, please mark the letter grade that is the closest equivalent to your grade average.)

(1) D or lower
(2) C-
(3) C
(4) B-
(5) B
(6) A-
(7) A

*P. Considering only your last two college years, approximately what *overall* grade average have you received? (If your college does not use letter grades, please mark the letter grade that is the closest equivalent to your grade average.)

(1) D or lower
(2) C-
(3) C
(4) B-
(5) B
(6) A-
(7) A

Q. Is there any geographic region in which you would *not* want to or could *not* attend graduate school? (Mark more than one response, if applicable.)

(1) Any region would be acceptable
(2) New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont)
(3) Mid-Atlantic (Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania)
(4) South (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia)
(5) Midwest (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin)
(6) Southwest (Arizona, California, Nevada, New Mexico, Oklahoma, Texas)
(7) West (Alaska, Colorado, Hawaii, Idaho, Montana, Oregon, Utah, Washington, Wyoming)

*R. About how many hours per week did you spend working for wages during your most recent school year?

(1) 0
(2) 1-5
(3) 6-10
(4) 11-20
(5) More than 20

*S. About how many hours per week did you spend in community (noncollege) activities during your most recent school year?

(1) 0
(2) 1-5
(3) 6-10
(4) 11-20
(5) More than 20

T. For which of the following achievements have you received an honor, award, prize, or other recognition? (Mark more than one if applicable.)

(1) Student government or organization
(2) Preprofessional—an award or prize for fieldwork or publication of a scholarly article or book
(3) Community service—election or appointment to a community service unit, activity, or group
(4) Literary—editing the college paper, yearbook, or literary magazine or having a poem, story, or article published in a public paper or magazine
(5) Artistic—a high rating in a music contest, a part in a play, opera, or show, or an award in an art competition
(7) Athletic—a letter in athletics
(8) None of the above categories

Major Field Code List

HUMANITIES
11 Archaeology
12 Architecture
26 Art History
13 Classical Languages
28 Comparative Literature
53 Dramatic Arts
14 English
29 Far Eastern Languages and Literature
15 Fine Arts, Art, Design
16 French
17 German
58 Italian
04 Linguistics
19 Music
57 Near Eastern Languages and Literature
20 Philosophy
21 Religious Studies or Religion
22 Russian
23 Spanish
24 Speech

SOCIAL SCIENCES
27 American Studies
81 Anthropology
82 Business and Commerce
83 Communications
84 Economics
85 Education (including M A in Teaching)
01 Educational Administration
09 Educational Psychology
70 Geography
92 Government
99 Guidance and Counseling
86 History
87 Industrial Relations and Personnel
88 International Relations
18 Journalism
89 Law
90 Library Science
91 Physical Education
92 Political Science
93 Psychology
94 Public Administration
55 Slavic Studies
79 Social Psychology
95 Social Work
96 Sociology
97 Urban Development (regional planning)

BIOLOGICAL SCIENCES
31 Agriculture
32 Anatomy
05 Audiology
33 Bacteriology
34 Biochemistry
35 Biology
36 Biophysics
37 Botany
38 Dentistry
39 Entomology
40 Forestry
06 Genetics
41 Home Economics
25 Hospital and Health Services Administration
42 Medicine
07 Microbiology
43 Nursing
77 Nutrition
44 Occupational Therapy
45 Optometry
46 Osteopathy
08 Parasitology
56 Pathology
03 Pharmacology
47 Pharmacy
48 Physical Therapy
49 Physiology
50 Public Health
51 Veterinary Medicine
52 Zoology

PHYSICAL SCIENCES
54 Applied Mathematics
61 Astronomy
62 Chemistry
78 Computer Sciences
63 Engineering, Aeronautical
64 Engineering, Chemical
65 Engineering, Civil
66 Engineering, Electrical
67 Engineering, Industrial
68 Engineering, Mechanical
69 Engineering, Other
71 Geology
72 Mathematics
73 Metallurgy
74 Mining
75 Oceanography
76 Physics
59 Statistics

*The asterisked items above were used in the stepwise regression analysis. The coding of the nominal item responses varied depending on the hypothesis being tested.

Appendix J

Intercorrelations, Means, and Standard Deviations for the Variables Reported in Table 4 and Table 5

	Group: Single (1) vs. Multiple (2) Repeater	2nd Testing minus 1st Testing: in Verbal Area	2nd Testing minus 1st Testing: in Quantitative Area	First Verbal Testing	Second Verbal Testing	First Quantitative Testing	Second Quantitative Testing	Citizen of United States (Yes=1, No=2)	How do you describe yourself?	Communicate better in English (Yes=1, No=2)	Number of full-time undergraduates in school of your Bachelor's degree	Best description of your undergraduate institution	Degree objective	Grade average in your undergraduate major	Overall average grade	Number of hours per week working for wages	Number of hours per week spent in noncollege activities in recent school years
	GRP	DV	DQ	V1	V2	Q1	Q2	B	D	E	F	G	J	O	P	R	S
GRP	1.00																
DV	- .07	1.00															
DQ	- .06	.05	1.00														
V1	- .13	- .11	.04	1.00													
V2	- .15	.36	.07	.89	1.00												
Q1	- .09	.09	- .20	.43	.44	1.00											
Q2	- .12	.11	.32	.44	.46	.86	1.00										
B	.04	- .01	.05	- .38	- .36	.06	.09	1.00									
D	.05	- .04	.01	- .29	- .30	- .21	- .20	.29	1.00								
E	.04	- .03	.04	- .34	- .34	.04	.07	.67	.31	1.00							
F	.04	- .01	- .01	- .01	- .02	.02	.01	- .04	- .04	- .01	1.00						
G	- .03	.02	.02	.15	.16	.06	.06	- .05	.00	- .04	- .55	1.00					
J	.04	.03	- .01	.28	.28	.23	.21	- .05	- .02	- .05	- .05	.06	1.00				
O	- .06	.06	.00	.29	.31	.24	.23	- .08	- .14	- .08	- .00	.03	.25	1.00			
P	- .02	.05	.00	.28	.29	.23	.23	- .06	- .15	- .06	- .00	.01	.28	.68	1.00		
R	.07	- .03	- .01	- .03	- .04	- .14	- .14	- .11	.07	- .07	.03	- .05	.09	- .04	- .01	1.00	
S	.03	- .02	.00	- .07	- .07	- .10	- .09	- .00	.09	.01	.03	- .02	.04	- .03	.00	.18	1.00
Mean	.10	25.75	21.86	451.49	477.24	478.26	500.12	1.11	.18	1.09	3.39	1.36	3.26	5.34	5.43	3.15	2.18
Standard Deviation	.30	56.98	64.87	114.58	122.31	121.99	126.34	.32	.39	.29	1.28	.48	1.05	1.08	1.04	1.56	1.07

Appendix K

Intercorrelations of Scores on Successive Retests for the Total Four-Administration Population

Test Administration	Verbal Scores (N=217)			
	1	2	3	4
2	.8591			
3	.8532	.8678		
4	.8267	.8420	.8633	
S.D.	113.7977	111.0935	106.0310	99.5342

Test Administration	Quantitative Scores (N=217)			
	1	2	3	4
2	.8211			
3	.8136	.8624		
4	.8096	.8410	.8602	
S.D.	121.0100	117.8918	111.8673	111.4044

Appendix L

Intercorrelations of Verbal Scores on Successive Retests for the Four-Administration Population, by Sex

Test Administration	Male (N=144)			
	1	2	3	4
2	.8565			
3	.8427	.8581		
4	.8257	.8417	.8618	
S.D.	112.2126	110.5748	105.7340	98.5127

Test Administration	Female (N=73)			
	1	2	3	4
2	.8550			
3	.8699*	.8840		
4	.8295	.8393	.8623	
S.D.	116.3298	111.0404	104.5958	99.8689

*Does not follow the simplex pattern.

Appendix M

Intercorrelations of Quantitative Scores on Successive Retests for the Four-Administration Population, by Sex

Test Administration	Male (N=144) 1	2	3	4
2	.8024			
3	.7819	.8510		
4	.7790	.8338	.8409	
S.D.	115.5759	114.8604	108.6614	108.1553

Test Administration	Female (N=73) 1	2	3	4
2	.8351			
3	.8513*	.8649		
4	.8417	.8321*	.8788	
S.D.	123.2027	116.2732	111.1621	109.6137

*Does not follow the simplex pattern.

GRE BOARD RESEARCH
REPORTS OF A TECHNICAL NATURE

Boldt, R. R. Comparison of a Bayesian and a Least Squares Method of Educational Prediction. GRE Board Professional Report GREB No. 70-3P, June 1975.

Campbell, J. T. and Belcher, L. B. Word Associations of Students at Predominantly White and Predominantly Black Colleges. GRE Board Professional Report GREB No. 71-6P, December 1975.

Carlson, A. B.; Reilly, R. R.; Mahoney, M. H.; and Casserly, P. L. The Development and Pilot Testing of Criterion Rating Scales. GRE Board Professional Report GREB No. 73-1P, October 1976.

Carlson, A. B.; Evans, F.R.; and Kuykendall, M. M. The Feasibility of Common Criterion Validity Studies of the GRE. GRE Board Professional Report GREB No. 71-1P, July 1974.

Donlon, T. F. An Exploratory Study of the Implications of Test Speededness. GRE Board Professional Report GREB No. 76-9P, March 1980.

Donlon, T. F.; Reilly, R. R.; and McKee, J. D. Development of a Test of Global vs. Articulated Thinking: The Figure Location Test. GRE Board Professional Report GREB No. 74-9P, June 1978.

Echternacht, G. Alternate Methods of Equating GRE Advanced Tests. GRE Board Professional Report GREB No. 69-2P, June 1974.

Echternacht, G. A Comparison of Various Item Option Weighting Schemes/A Note on the Variances of Empirically Derived Option Scoring Weights. GRE Board Professional Report GREB No. 71-17P, February 1975.

Echternacht, G. A Quick Method for Determining Test Bias. GRE Board Professional Report GREB No. 70-8P, July 1974.

Evans, F. R. The GRE-Q Coaching/Instruction Study. GRE Board Professional Report GREB No. 71-5aP, September 1977.

Fredericksen, N. and Ward, W. C. Development of Measures for the Study of Creativity. GRE Board Professional Report GREB No. 72-2P, June 1975.

Levine, M. V. and Drasgow, F. Appropriateness Measurement with Aptitude Test Data and Estimated Parameters. GRE Board Professional Report GREB No. 75-3P, March 1980.

McPeek, M.; Altman, R. A.; Wallmark, M.; and Wingersky, B. C. An Investigation of the Feasibility of Obtaining Additional Subscores on the GRE Advanced Psychology Test. GRE Board Professional Report GREB No. 74-4P, April 1976.

Powers, D. E.; Swinton, S.; Thayer, D.; and Yates, A. A Factor Analytic Investigation of Seven Experimental Analytical Item Types. GRE Board Professional Report GREB No. 77-1P, June 1978.

Powers, D. E.; Swinton, S. S.; and Carlson, A. B. A Factor Analytic Study of the GRE Aptitude Test. GRE Board Professional Report GREB No. 75-11P, September 1977.

Reilly, R. R. and Jackson, R. Effects of Empirical Option Weighting on Reliability and Validity of the GRE. GRE Board Professional Report GREB No. 71-9P, July 1974.

Reilly, R. R. Factors in Graduate Student Performance. GRE Board Professional Report GREB No. 71-2P, July 1974.

Rock, D. A. The Identification of Population Moderators and Their Effect on the Prediction of Doctorate Attainment. GRE Board Professional Report GREB No. 69-6bP, February 1975.

Rock, D. A. The "Test Chooser": A Different Approach to a Prediction Weighting Scheme. GRE Board Professional Report GREB No. 70-2P, November 1974.

Sharon, A. T. Test of English as a Foreign Language as a Moderator of Graduate Record Examinations Scores in the Prediction of Foreign Students' Grades in Graduate School. GRE Board Professional Report GREB No. 70-1P, June 1974.

Swinton, S. S. and Powers, D. E. A Factor Analytic Study of the Restructured GRE Aptitude Test. GRE Board Professional Report GREB No. 77-6P, February 1980.

Ward, W. C.; Frederiksen, N.; and Carlson, S. B. Construct Validity of Free-Response and Machine-Scorable Versions of a Test of Scientific Thinking. GRE Board Professional Report GREB No. 74-8P, November 1978.

Ward, W. C. and Frederiksen, N. A Study of the Predictive Validity of the Tests of Scientific Thinking. GRE Board Professional Report GREB No. 74-6P, October 1977.